What is claimed is:

1. A phoneme dividing method using a multilevel neural network applied to a phoneme dividing apparatus having a voice input portion for outputting a vocal sample digitally converted from voice made, a preprocessor for extracting a characteristic vector suitable for phoneme division, from the vocal sample input from the voice input portion, a multi-layer perceptron (MLP) phoneme dividing portion for finding and outputting the border of phoneme, using the characteristic vector of the preprocessor, and a phoneme border outputting portion for outputting position information on the border of phoneme of the MLP phoneme dividing portion in the form of frame position, said method comprising the steps of:

(a) sequentially segmenting and framing voice with digitalized voice samples, extracting characteristic vectors by vocal frames, and extracting an inter-frame characteristic vector of the difference between nearby frames of the characteristic vectors by frames, to thereby normalize the maximum and minimum of said characteristics;

(b) initializing weights present between an input layer and hidden layer and between the hidden layer and output layer of said MLP, designating an output target data of said MLP, inputting said characteristic vectors to said MLP for learning, and storing and finishing information on the weight obtained through learning and the standard of said MLP if the reduction

15

rate of mean squared error converges within a permissible limit; and

(c) reading the weight obtained in said step (b), receiving said characteristic vectors, performing an operation of phoneme border discrimination to generate an output value, discriminating the phoneme border according to the output value, and if the current analyzed frame arrives two frames preceding the final frame of incoming voice, outputting a frame number indicative of the border of phoneme as a final result.

2. The method as claimed in claim 1, wherein the voice framing of said step (a) is performed by taking a Hamming window in a length of 16 msec every 10 msec, with respect to the overall incoming vocal samples.

3. The method as claimed in claim 1, wherein the phoneme border discrimination of said step (c) is performed in such a manner that output values generated through operation are compared, and then it is determined that if output value OUT(0) is positive, an analyzed frame is the border of phonemes, and if output value OUT(1) is positive, the frame is not the border of phonemes.